TRAFFIC MANAGEMENT

Phase 2 project

Abstract:

The project involves integrating historical traffic data and machine learning algorithms to predict congestion patterns.

This proposal advocates for the integration of historical traffic data and machine learning algorithms to predict traffic congestion patterns, thus enhancing traffic management. The approach involves collecting and preprocessing traffic data, extracting relevant features, selecting appropriate machine learning models, and training them using historical data.

Real-time data integration allows for continuous learning and dynamic predictions, facilitating timely alerts and optimized traffic management strategies. The system's effectiveness is evaluated through real-world deployment, providing insights into its potential to significantly improve traffic flow and reduce congestion for commuters.

1.Data Collection: Gather historical traffic data, including traffic volume, speed, and congestion incidents. Sources can include traffic cameras, GPS data from vehicles, and public transportation data.

3 types of data collection

- **Equipment Data**

This is the type of data that concerns the status of the IoT devices. Equipment data is collected in real time to facilitate **activities of predictive maintenance**.

- **Submeter Data**

Submetering allows property owners to **automate the measurement** of individual utility usage in multi-user settings.

Submeter data can be **collected in buildings** where multiple tenants use resources like water, electricity, gas, or cable

Materials and Methods:

A Road Traffic Prediction Dataset from Huawei Munich Research Center is used, which is a public dataset for traffic prediction derived from a variety of traffic sensors, i.e., induction loops , it is important to note that, at present, there are a few public datasets . The data can be used to forecast traffic patterns and modify stop-light control parameters. The dataset contains recorded data from six crosses in the urban area for 56 days, in the form of flow time series, depicting the number of vehicles passing every five minutes for a whole day, which is recommended for short-term predictions For this research, four of the six intersections are used to simulate four lanes of an intersection.

2.**Data Preprocessing**: Clean and preprocess the data to handle missing values, outliers, and noise. This step is crucial for ensuring the quality of input data for machine learning models.

Data processing cycle:

- ***Input***

Input is the first stage of the data processing cycle. It is a stage in which the collected data is converted into a machine-readable form so that a computer can process it. This is a very important stage since the data processing output is completely dependent on the input data (“garbage in - garbage out”).

- ***Processing***

In the processing stage, a computer transforms the raw data into information. The transformation is carried out by using different data manipulation techniques, such as:

- Classification: Data is classified into different groups.
- Sorting: Data is arranged in some kind of an order (e.g. alphabetical).
- Calculation: Arithmetic and logical operations are performed on numeric data.

Output:

In the last stage, output is received. This is the stage where the processed data is converted into human-readable form and presented to the end user as useful information. Also, the output of data processing can be stored for future use.

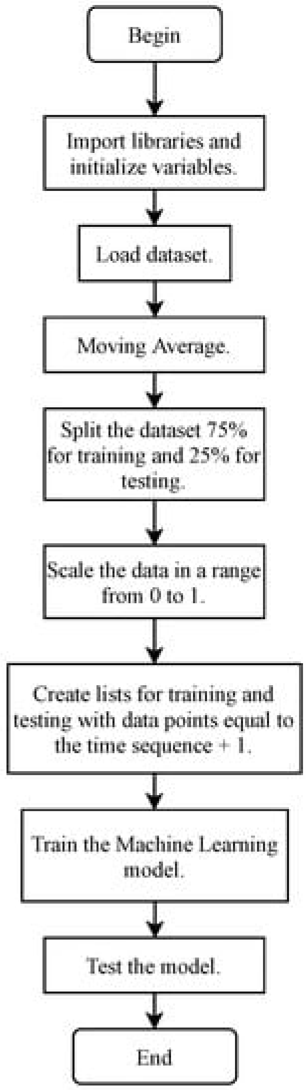
Feature Engineering: Extract relevant features from the data, such as time of day, day of the week, weather conditions, and special events. These features can help improve the accuracy of congestion predictions.

Machine Learning Model Selection: Choose appropriate machine learning algorithms for traffic prediction. Common choices include regression models, decision trees, random forests, or more advanced techniques like neural networks.

Training and Validation: Split the data into training and validation sets to train and evaluate the chosen machine learning model. Use metrics like Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE) to assess model performance.

Machine Learning Methods:

Five regression models from the scikit-learn library in Python are used: Linear Regression, Gradient Boosting Regressor, MultiLayer Perceptron Regressor, Stochastic Gradient Descendent Regressor and Random Forest Regressor, all of them with default parameters and a random state equal to zero for the reproducibility of the experiment. A reshape of 'X' for both training and testing is made because these models require a 2D array instead of the 3D used in the RNN's; after that, the models are fed with the training split. The methodology summary presented in the form of a flowchart.



Real-Time Data Integration: Implement a system to integrate real-time traffic data into the model for continuous learning and updates. This allows the model to adapt to changing traffic conditions.

Prediction and Alerting: Use the trained model to make real-time predictions of traffic congestion patterns. Set up alerting systems to notify traffic management authorities and drivers about potential congestion ahead.

Feedback Loop:
Continuously collect new data and use it to retrain the machine learning model periodically. This helps the model stay up-to-date with evolving traffic patterns.

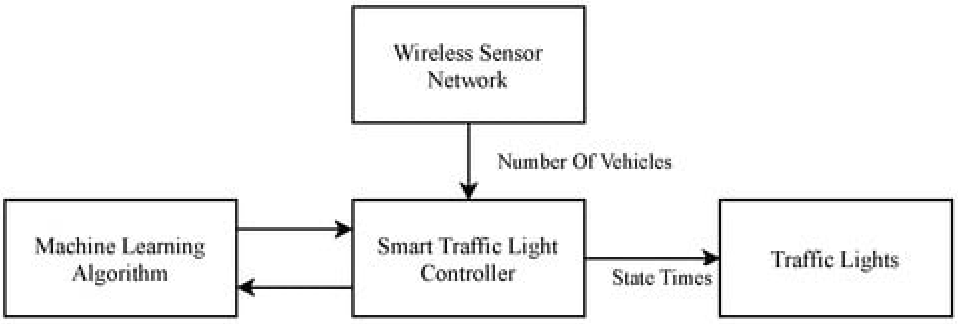
Deployment:
Deploy the system in relevant traffic management operations, such as adjusting traffic signals, rerouting traffic, or providing information to drivers via mobile apps or electronic signs.

Monitoring and Optimization:
Regularly monitor the system's performance and make improvements as needed. Optimization may involve fine-tuning model hyperparameters or adding new features.

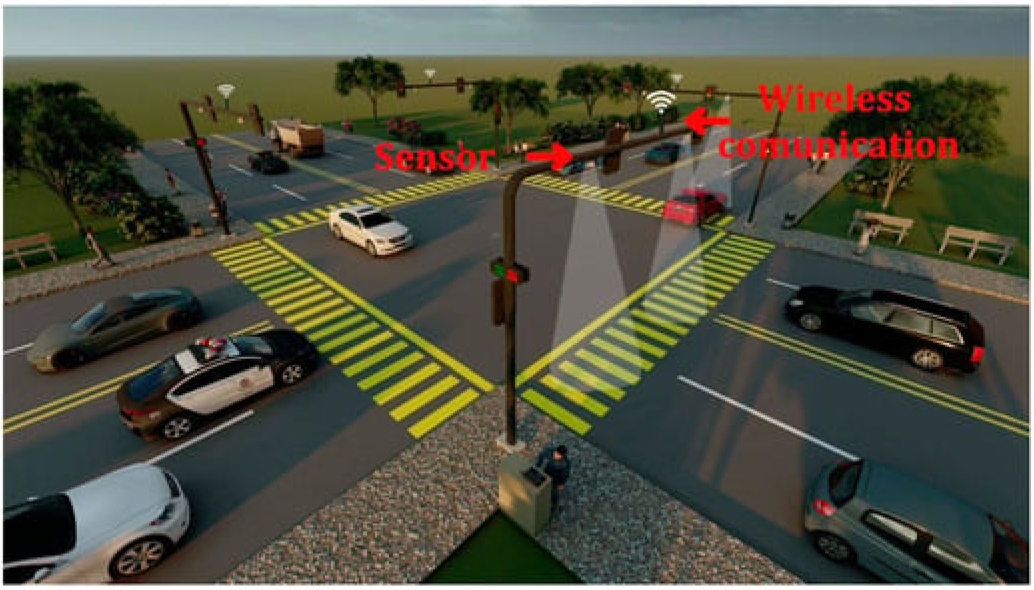
Proposed Usage Scenario:

These models can be used in a smart traffic light controller, fed by traffic sensors that count the number of vehicles passing through a lane every certain period; with these readings, a database similar to the one used in this paper can be created. Once the database is generated, the **ML** model can be trained for each intersection. Then the traffic flow for the next period can be predicted by using a given number of past readings.

Once the prediction is made, it will allow better programming of the times of each state, either manually by an operator or automatically using an algorithm to calculate the optimal times of the traffic light states. The whole process can be carried out by wirelessly communicating the traffic light with a central station or at the controller itself fig a shows a block diagram of the main elements of the proposed system and fig b its representation in a real-world scenario.



(a)



(b)

Conclusion:

The proposed traffic prediction and congestion detection system amalgamate historical traffic data and machine learning methodologies to offer a comprehensive approach toward effective traffic management. Beginning with the collection and preprocessing of historical traffic data, the system cleans and structures the data to prepare it for analysis. Feature engineering identifies relevant variables impacting traffic congestion, aiding in the creation of predictive models.

A pivotal step is the selection and training of machine learning models, allowing the system to learn patterns and trends from the historical data. The performance of these models is evaluated to ensure accuracy and reliability. Real-time traffic data integration enhances predictions, providing up-to-date insights into traffic conditions.

Visualizations and reports serve as intuitive interfaces, enabling stakeholders to interpret predicted congestion patterns and make informed decisions. Continuous feedback from traffic managers and users ensures the system's adaptability and efficiency. Iterative refinement of models based on feedback culminates in a dynamic and evolving system, proficient in optimizing traffic flow and reducing congestion.

In essence, this proposed approach offers a promising solution to the complex challenges of traffic management, leveraging technology and data-driven insights to pave the way for improved urban mobility and efficient transportation systems.